

ARTICLES

Speciation through sensory drive in cichlid fish

Ole Seehausen^{1,2}, Yohey Terai³, Isabel S. Magalhaes^{1,2}, Karen L. Carleton⁴, Hillary D. J. Mrosso⁵, Ryutaro Miyagi³, Inke van der Sluijs^{6,†}, Maria V. Schneider^{2,†}, Martine E. Maan^{6,†}, Hidenori Tachida⁷, Hiroo Imai⁸ & Norihiro Okada³

Theoretically, divergent selection on sensory systems can cause speciation through sensory drive. However, empirical evidence is rare and incomplete. Here we demonstrate sensory drive speciation within island populations of cichlid fish. We identify the ecological and molecular basis of divergent evolution in the cichlid visual system, demonstrate associated divergence in male colouration and female preferences, and show subsequent differentiation at neutral loci, indicating reproductive isolation. Evidence is replicated in several pairs of sympatric populations and species. Variation in the slope of the environmental gradients explains variation in the progress towards speciation: speciation occurs on all but the steepest gradients. This is the most complete demonstration so far of speciation through sensory drive without geographical isolation. Our results also provide a mechanistic explanation for the collapse of cichlid fish species diversity during the anthropogenic eutrophication of Lake Victoria.

The sensory drive hypothesis for speciation^{1,2} predicts that adaptation in sensory and signalling systems to different environments in allopatry may cause premating isolation on secondary contact of populations. Recent theoretical work suggested that sensory drive can lead to the evolution of colour polymorphisms^{3,4} and speciation⁵, even in the absence of geographical isolation, when the light environment is heterogeneous. However, the only case of sympatric sister species, in which assortative mating has been shown to be facilitated by sensory drive, were sticklebacks in British Columbia⁶. Here we provide ecological, population genetic and molecular evidence for each of the predictions of sensory drive speciation² in sympatric cichlid fish inhabiting light gradients in Lake Victoria (East Africa).

Lake Victoria is spatially highly heterogeneous in water clarity and ambient light^{7,8}, and there is much evidence that the cichlid visual system has been under strong diversifying selection during the adaptive radiation of cichlids into several hundred species in Lake Victoria⁹. Vertebrate visual pigments consist of a light-absorbing component, the chromophore, and a protein moiety, the opsin¹⁰. Spectral sensitivity is determined by the chromophore (A1 or A2 pigments), and by its interaction with the amino acid residues lining the retinal-binding pocket of the opsin in which the chromophore lies¹¹. Eight different visual pigments have been found in all haplochromine cichlids^{12–14}, but only a subset of these is expressed in any individual species^{12,14,15}. Several *Pundamilia* species from Lake Victoria expressed the same complement of four opsin genes: short-wavelength-sensitive opsin gene 2a (*SWS2A*, $\lambda_{\max} \sim 455$ nm) in single cones; rhodopsin-like (*RH2*, $\lambda_{\max} \sim 528$ nm) and long-wavelength-sensitive opsin gene (*LWS*, $\lambda_{\max} \sim 565$ nm) in double cones; and rhodopsin (*RH1*, $\lambda_{\max} \sim 505$ nm) in rods¹⁶. Of these, the *LWS* opsin gene is by far the most variable among Lake Victoria cichlids^{13,17}, with sequence variation being five times greater than

in Lake Malawi cichlids despite a tenfold greater age of the latter species flock¹⁸.

Female Lake Victoria cichlids have mating preferences for conspicuously coloured males¹⁹. Perception of conspicuousness is influenced by ambient and background light, signal transmission, receiver sensitivity and higher level processing²⁰. Sympatric pairs of closely related cichlid species, one with red and one with blue nuptial colouration (Fig. 1 and Supplementary Fig. 3), are common in Lake Victoria⁸. Visual pigments have been compared for two pairs, and behavioural light detection thresholds measured in three. In each pair, the red species has its *LWS* λ_{\max} at a longer wavelength^{16,21}, with a lower detection threshold for red but a higher one for blue light^{22,23}. These observations are consistent with a role for sensory drive in speciation, whereby interaction between ambient light, natural-selection-driven divergence of visual sensitivities and sexual selection for conspicuous male colours leads to the fixation of different male colours^{1,2,16,23}.

Examining the role of environmental gradients in speciation requires tests to replicate gradients, as is recognized both in evolutionary ecology^{24–26} and in population genomics²⁷. A recent model of clinal speciation through sensory drive⁵, as well as other models of clinal speciation^{28–30}, predicts the greatest probability of speciation on gradients of intermediate slope. There, migration rates are sufficiently low to be compensated for by selection, but are sufficiently high to generate significant migration load³¹ and intermediate genotypes with a poor fit to the local environment. Migration load and reduced fitness of intermediate genotypes lead to disruptive selection, which may be required for the evolution of assortative mating through reinforcement-like mechanisms^{28–30}. Previously we demonstrated adaptive evolution in the *LWS* opsin gene of the Lake Victoria cichlid fish *Neochromis greenwoodi* and *Mbipia mbipi* along very shallow gradients of light colour mediated by variation in turbidity

¹Institute of Zoology, University of Bern, Baltzerstr. 6, CH-3012 Bern, Switzerland. ²Eawag, Swiss Federal Institute for Aquatic Science and Technology, Centre of Ecology, Evolution & Biogeochemistry, Department of Fish Ecology & Evolution, 6047 Kastanienbaum, Switzerland. ³Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan. ⁴Department of Biology, University of Maryland, College Park, Maryland 20742, USA. ⁵Tanzania Fisheries Research Institute, Mwanza Centre, PO Box 475 Mwanza, Tanzania. ⁶Department of Animal Ecology, Institute of Biology, Leiden University, PO Box 9516, 2300 RA Leiden, The Netherlands. ⁷Department of Biology, Faculty of Sciences, Kyushu University, Ropponmatsu, Fukuoka 810-8560, Japan. ⁸Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, 484-8506 Japan. [†]Present addresses: Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montréal, Québec H3A 1B1, Canada (I.v.d.S.); The European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK (M.V.S.); University of Texas at Austin, Integrative Biology, 1 University Station C0930, Austin, Texas 78712, USA (M.E.M.).

between islands⁹. *LWS* genotype frequencies and male colour morph frequencies formed correlated clines, but, even though populations at opposite ends of one gradient fixed different *LWS* alleles, all populations retained polymorphism for colour, indicating that speciation remained incomplete⁹.

Here we investigate populations of cichlid fish living on light gradients primarily mediated by water depth within islands in Lake Victoria. *Pundamilia pundamilia* and *Pundamilia nyererei*²² (Fig. 1a and Supplementary Fig. 3) are geographically fully sympatric. Within islands, they have narrowly parapatric depth ranges. Where they are phenotypically well differentiated, *P. pundamilia* has blue–grey male nuptial colouration whereas *P. nyererei* nuptial males are yellow with a bright crimson-red dorsum. Females of both are cryptically yellowish and have mating preferences for the nuptial colouration of conspecific males^{33,34}. The red *P. nyererei* occurs at greater mean water depths, in more red-shifted ambient light than the blue *P. pundamilia*²³. *P. nyererei* have a lower threshold for the detection of red light, whereas *P. pundamilia* possess a lower threshold for detection of blue light²³. Earlier we found that red and blue fish tended to possess different alleles at the *LWS* opsin gene locus¹⁶. Here we fully develop this system to test predictions of sensory drive speciation.

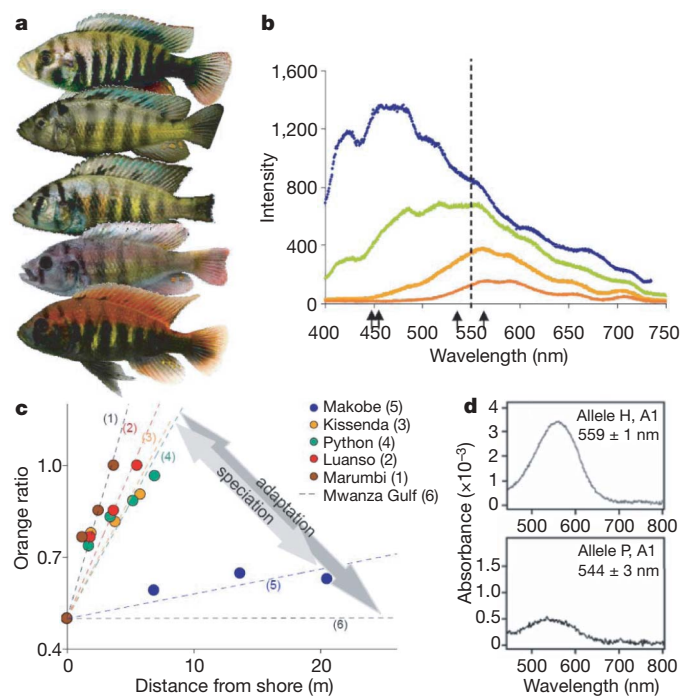


Figure 1 | Male phenotypes, light gradients and *LWS* opsin absorbance. **a**, Variation in male nuptial colouration. Five phenotype classes from 0 ('blue', typical *P. pundamilia*; top) to 4 ('red', typical *P. nyererei*; bottom). **b**, An example of a moderately steep light gradient (Python island): surface light spectrum (blue) and three subsurface light spectra measured at 0.5 m (green), 1.5 m (orange) and 2.5 m (red) water depth. The line through 550 nm indicates the divide used to calculate the transmittance orange ratio. Arrows indicate peak absorbance of two opsin pigments: main allele groups at *LWS* opsin locus (544 nm and 559 nm) and known range of peak absorbance at *SWS2A* locus¹⁶. **c**, Slopes of seven different light gradients. The lines for two shallow gradients overlay each other and are together labelled 'Mwanza Gulf'. For this line, the x-axis represents the distance from clear water (rather than from shore). Significant differentiation in opsin genes was observed on all gradients with slopes equal to or shallower than the Kissenda (orange) line, but speciation was observed only on gradients with slopes between the Kissenda (orange) and the Makobe (blue) lines. The dark grey arrow indicates region with divergent adaption at *LWS* opsin gene, and the light grey arrow indicates region with speciation. **d**, Absorption spectra of *LWS* pigments evaluated by the dark–light difference spectra⁹. The *LWS* pigments were reconstituted from the H allele with A1 retinal (top) and from the P allele with A1 retinal (bottom). λ_{\max} values (with standard errors) are indicated.

If sensory drive caused speciation into a red and a blue species, we expected to find: (1) variation in the *LWS* opsin sequence at amino acid positions where they shift λ_{\max} ; (2) an association of such sequence variation with water depth, such that more red-shifted alleles occur at greater depth; and (3) an association of *LWS* alleles with the predominant male nuptial colouration of a population, such that populations with predominantly red-shifted opsin alleles have predominantly red males. Furthermore, if disruptive selection was required to complete speciation through the evolution of assortative mating, we predicted that the strongest associations between *LWS* alleles, water depth and colour occur on intermediate light slopes (prediction (4)). For testing prediction (4), we compared the data from the depth-mediated gradients of this study with data we had collected earlier on populations occupying the same depth at different islands with different turbidities⁹ (see Supplementary Information).

Light, depth and colour

We examined depth-mediated light gradients at five islands. The light climate of Lake Victoria is dominated by effects of particulate (non-phytoplankton) matter, selectively absorbing and scattering light of short wavelengths³⁵, causing successive shifts of ambient light towards longer wavelengths with increasing water depth (this study), and also with increasing turbidity (earlier study)^{7,8}. The rate at which ambient light changes with increasing depth is positively correlated with turbidity⁸ (difference between islands in this study). The cichlids we study feed and breed right above and within the rocky substrate. We characterize depth-associated light gradients in their habitat by the change in the 'transmittance orange ratio' that occurs per metre as one moves outwards from the shore into the lake along the lake floor (the 'light slope', see Methods and Fig. 1b). Steeper slopes occur with more turbid water and steeper shores (Table 1). The steepest light slopes occurred at the most turbid sites, Marumbi and Luanso islands (Table 1 and Fig. 1c). Intermediate slopes occurred at Kissenda and Python islands, and the shallowest slope at Makobe island. The latter was still steeper than all the turbidity-mediated light slopes of our earlier work⁹. The size of the light differential between the ends of the gradients was similar between the five depth-mediated gradients, and larger than on the turbidity-mediated gradients (Table 1 and Supplementary Table 1).

Mapping the microdistribution of phenotypes on the five depth-mediated gradients using data from 960 males (Fig. 2a) revealed significant differences between islands. It showed the absence of any association between colour and ambient light (water depth) at Marumbi and Luanso (analysis of variance, ANOVA: $df = 2$, $F = 1.1$, $P = 0.3$, and $df = 2$, $F = 0.3$, $P = 0.7$, respectively), but significant associations at all other sites (ANOVA: $df = 2$ (Kissenda), $df = 1$ (Python, Makobe), $F > 50$, $P < 0.0001$), and increasing strength of association with decreasing light slope (F ratio against slope, logarithmic regression, $df = 4$, $R^2 = 0.87$, $P = 0.021$; Fig. 3). Blue phenotypes are associated with shallow waters (<3 m) in all locations, whereas red phenotypes occur in shallow waters only on the steepest gradients, and become restricted to greater depths with decreasing light slope. Frequency distributions of male nuptial colour phenotypes differ significantly between islands too (Fig. 2b). Distributions are unimodal and skewed towards blue on the two steepest gradients. They are bimodal with few intermediates on gradients of intermediate slope, and consist of two discrete classes, blue and red, on the shallowest within-island gradient.

Table 1 | The five environmental gradients of this study

Island	Water clarity (cm Secchi) (mean \pm s.d.)	Shoreline slope (mean \pm s.d.)	Light slope	Light differential
Marumbi island	53 \pm 8	0.82 \pm 0.15	1.4×10^{-1}	0.50
Luanso island	50 \pm 10	0.54 \pm 0.05	9.6×10^{-2}	0.50
Kissenda island	78 \pm 24	0.52 \pm 0.12	7.9×10^{-2}	0.50
Python island	96 \pm 21	0.58 \pm 0.24	7.6×10^{-2}	0.50
Makobe island	225 \pm 67	0.15 \pm 0.04	8×10^{-3}	0.35

LWS gene variation, light and colour

We observed 13 polymorphic sites (3 synonymous, 10 nonsynonymous) among the *LWS* sequences (Supplementary Table 6). Three nonsynonymous substitutions occurred at high frequencies. From

the bovine rhodopsin crystal structure³⁶ we inferred that two of these variable amino acid positions, 216 (nucleotide site 647) and 275 (823 and 824), are located in or near the retinal-binding pocket. The third one was position 230 (688), one of the tuning sites of human red/

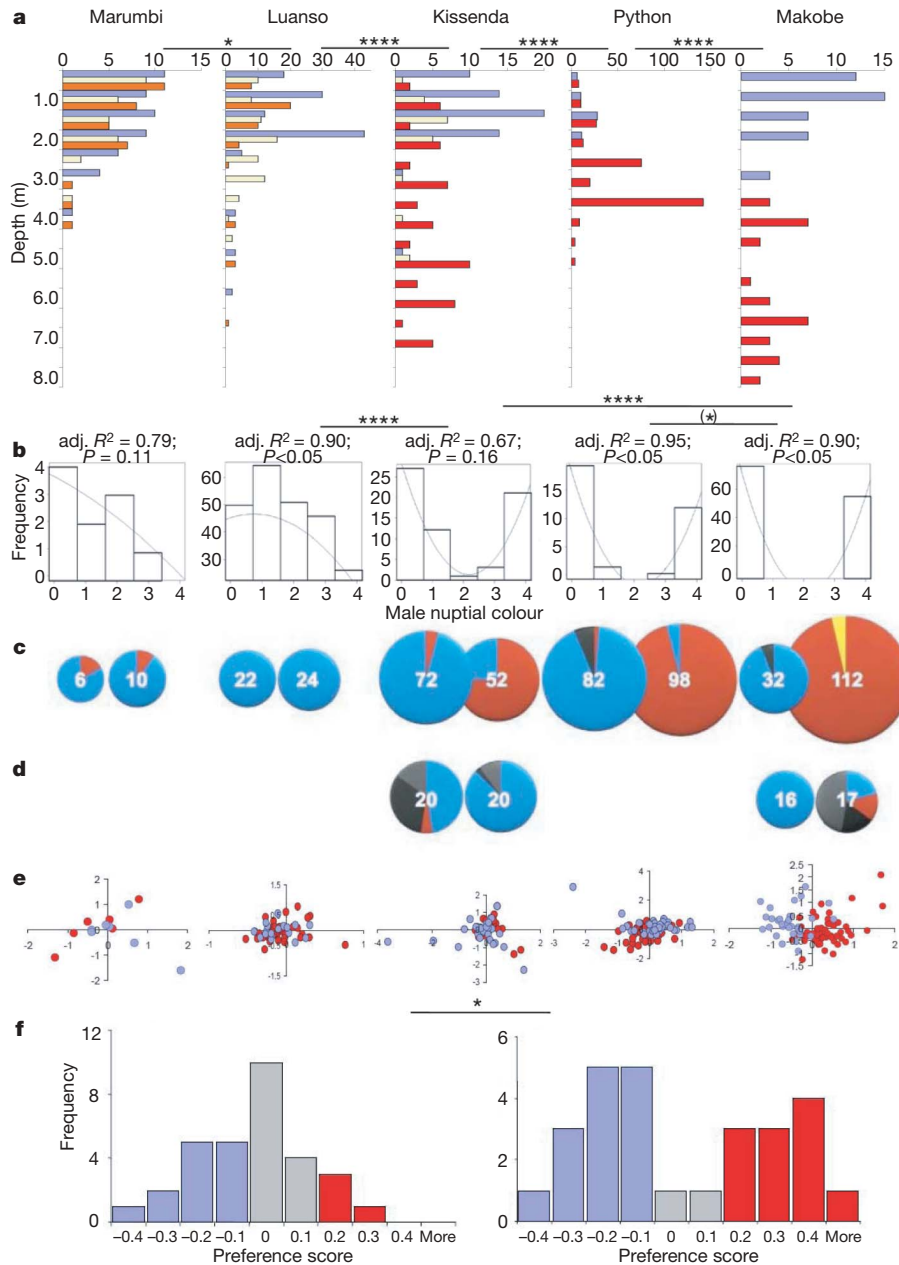


Figure 2 | Ecological, phenotypic, genetic and behavioural differentiation between blue and red *Pundamilia* nuptial phenotypes at five islands. All data for the same island are presented in the same column. Significant differences between islands indicated by asterisks (all tests two-tailed): * $P < 0.05$, **** $P < 0.0001$, (*) $P < 0.1$. **a**, Depth distributions of male nuptial colour phenotypes. Blue bars, blue; pale yellow bars, intermediate; and orange or red bars, red (orange if dominated by class 3; red if dominated by class 4). Significance levels of differences between islands in the divergence between red and blue reported as P values of G -tests. **b**, Frequency distributions of male nuptial colour phenotypes (see Fig. 1a and text). Lines are quadratic fits; R^2 and significance levels indicated. Significance levels of differences between islands reported as P values of G -tests. **c**, Frequencies of functional allele groups at the *LWS* opsin gene by island and male colour (left, blue; right, red). Numbers report sample sizes of completely sequenced haplotypes. For Marumbi and Luanso islands, only the haplotypes of those individuals are included that could be assigned to 'blueish' and 'reddish' phenotypes (altogether 24 and 54 haplotypes were sequenced from Marumbi and Luanso,

respectively). Fish from Marumbi were divided into classes 0 + 1 and classes 2 + 3. Fish from Luanso were divided into classes 0 + 1 and 2–4. At all other islands, only fish of phenotype classes 0 and 4 were included. Alleles of the P group shown in blue, alleles of the H group in red, M3 alleles in yellow, and other alleles in grey. **d**, Allele frequencies at the *SWS2A* opsin gene and nuptial colour class. The *SWS2A* P allele shown in blue, the N allele in red, other alleles in black, and alleles not determined in grey. **e**, Individuals plotted on first and second axes of a factorial correspondence analysis of genetic variance calculated from 11 unlinked microsatellite loci. Colours indicate pooled male nuptial colour classes as described in **c**. **f**, Histograms of female mating preferences at Luanso island⁴¹ (left) and Python island⁴⁰ (right, includes new data). Blue, preference classes in which most females had statistically significant individual preferences for blue males; red, preference classes in which most females had significant preferences for red males; grey, preference classes in which females had no significant mating preference. Significance level of the difference in the frequency distributions between the two islands reported as P value of a G -test.

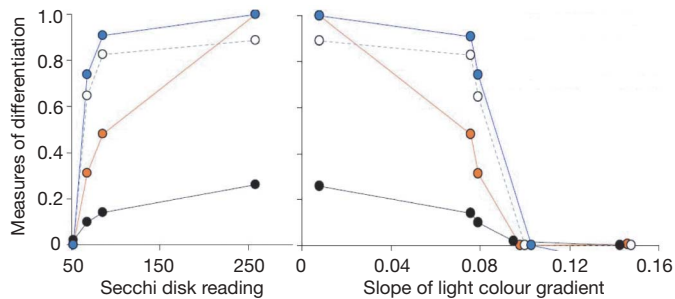


Figure 3 | Measures of differentiation between sympatric *Pundamilia* phenotypes plotted against water transparency (left) and light slope (right). Blue symbols and line: Spearman rank correlations between colour and *LWS* genotype (best fit to water clarity $R^2 = 0.79$, $P = 0.045$, $df = 4$; best fit to light slope $R^2 = 0.69$, $P(\text{one-tailed}) = 0.042$, $df = 4$). Open symbols and dashed line: *LWS* F_{ST} between red and blue phenotypes (best fit to water clarity, $R^2 = 0.79$, $P = 0.044$, $df = 4$; best fit to light slope, $R^2 = 0.65$, $P(\text{one-tailed}) = 0.045$, $df = 4$). Filled orange symbols and orange line: association between colour and water depth (ANOVA F ratios (normalized to range 0–1 for display); best fit to water clarity $R^2 = 0.99$, $P = 0.000$; best fit to light slope $R^2 = 0.87$, $P = 0.021$; both $df = 4$). Filled black symbols and black line: microsatellite F_{ST} (multiplied by 10 for display) between red and blue phenotypes (best fit to water clarity, $R^2 = 0.99$, $P = 0.000$; best fit to light slope, $R^2 = 0.90$, $P(\text{one-tailed}) = 0.013$; both $df = 4$).

green opsin³⁷. Focusing on these three positions, we divided alleles into three groups described previously⁹: the H group (all alleles with 216Y, 230A, 275C), the P group (216F, 230T, 275I) and the M3 group (216Y, 230T, 275I). M3 alleles can be considered recombinants or intermediate between H and P alleles. H and P alleles differed in only the 3 amino acid positions 216, 230 and 275. Substitutions at the other 7 nonsynonymous sites were rare and resulted in other allele variants (Supplementary Table 5).

We reconstituted the *LWS* pigments from P alleles *in vitro* with A1-derived retinal, and measured their absorption spectra, as previously done for the H alleles⁹ (Fig. 1d). The peak spectral sensitivity (λ_{max}) of the A1 pigment of the P allele was blue-shifted by 15 nm relative to the H allele. The λ_{max} values of cone outer segments expressing either P or H pigments were measured previously by microspectrophotometry, reporting too that P pigments were blue-shifted relative to H pigments¹⁶. Hence, the absorption spectra of P and H alleles seem to be adapted to shallower and deeper water light environments in Lake Victoria, respectively (Fig. 1b, d), supporting prediction (1).

Light gradients with slopes steeper than 0.09 were inhabited by populations with one or two different *LWS* alleles, whereas up to six different alleles were present on less steep gradients (Table 1 and Supplementary Tables 2 and 6). On these gradients of steepness $0.008 \leq x \leq 0.09$, H alleles were strongly associated with red nuptial

colouration ($\chi^2 > 66$, $df = 1$, $P < 0.0001$; Spearman correlation coefficients 0.74, 0.91 and 1, respectively, for slopes 0.079, 0.076 and 0.008; $P < 0.0001$), and were rare in blue phenotypes (Fig. 2c, Supplementary Table 6 and Supplementary Information), supporting prediction (3).

A strong association between *LWS* alleles and water depth emerges from these results, supporting prediction (2): at Marumbi and Luanso islands, most individuals reside in waters less than 3 m deep. P alleles strongly dominate. At all other islands, only the blue phenotype is confined to depths less than 3 m, and P alleles predominate among these fish, even where gene exchange with the red phenotype is frequent (see later). The sweep to high frequency of H alleles in the red phenotype is associated with shifting larger fractions of the population to depths beyond 3 m. At Kissenda island, 75% of the *LWS* alleles of the red population belong to the red-shifted H group. The proportion of H alleles, residing in red individuals, increases to Python island and further to Makobe island, associated with successively increasing fractions of the red population living in deep water (Fig. 2a versus 2c). Red and blue phenotypes were highly significantly differentiated at the *LWS* locus at Kissenda, Python and Makobe islands (F_{ST} (fixation index) 0.65, 0.83, 0.89), but neither at Luanso nor at Marumbi islands (F_{ST} 0.00).

Gene flow at neutral loci

A sensory drive model of speciation predicts that the rate of divergence at the opsin loci should exceed the rate of divergence at neutral loci. Our data are fully consistent with this prediction (Table 2). Pairwise F_{ST} between sympatric blue and red phenotypes estimated from 11 microsatellite loci reveal no differentiation at Marumbi or Luanso islands (Fig. 2e), consistent with the unimodal frequency distributions of male nuptial colour variants and the absence (Marumbi) or rarity of really red males. Pairwise F_{ST} at all other islands suggest significant, albeit weak, differentiation, consistent with the strongly bimodal frequency distributions of male nuptial colour variants and the emergence of the really red phenotype at those islands. Whereas F_{ST} at the *LWS* locus jumps from 0 at Marumbi and Luanso to 0.65 at Kissenda, F_{ST} at microsatellite loci increases gradually and much more slowly (Figs 2c, e and 3). The number of microsatellite loci carrying the signature of differentiation increases steadily from Marumbi and Luanso (0 out of 11) to Makobe island (7 out of 11; Table 2), consistent with the successive disappearance of intermediate phenotypes.

With the exception of Makobe island, all microsatellite F_{ST} among sympatric red and blue phenotypes are smaller than F_{ST} between any two allopatric populations of the blue phenotype, and 7 out of 10 of the red phenotype (Supplementary Fig. 1a and Supplementary Table 3). Even the largest between-phenotype F_{ST} at Makobe is smaller than most within-phenotype F_{ST} between islands. This suggests either more

Table 2 | Pairwise F_{ST} statistics between sympatric phenotypes

Island	Marumbi island	Luanso island	Kissenda island	Python island	Makobe island
Light slope	0.144	0.096	0.079	0.076	0.008
F_{ST} at <i>LWS</i> opsin locus	0.000	0.000	0.648	0.826	0.890
F_{ST} at microsatellite loci					
Ppun21	0.000	0.000	0.010	0.006	0.023
Ppun7	0.000	0.000	0.003	0.023	0.013
Ppun5	0.000	0.002	0.002	0.000	0.010
Ppun32	0.041	0.005	0.000	0.016	0.080
Ppun17	0.000	0.000	0.000	0.046	0.027
OSU16d	0.017	0.002	0.011	0.006	0.020
OSU20d	0.002	0.000	0.040	0.004	0.008
OSU19t	0.000	0.022	0.013	0.013	0.032
TMO5	0.000	0.013	0.000	0.012	0.010
Pzeb3	0.000	0.000	0.002	0.048	0.107
Pzeb5	0.000	0.000	0.024	0.024	0.049
Multilocus (11 μ sats)	0.000	0.002	0.010	0.014	0.026

Significant F_{ST} ($P < 0.05$) are shown in bold.

gene flow or more recent divergence between phenotypes within islands than between island populations of the same phenotype. It implies either parallel maintenance of phenotypic differentiation in the face of gene flow, or parallel sympatric speciation. All H alleles as well as the most frequent P allele are shared with several distantly related cichlid species (Supplementary Fig. 4). The two *Pundamilia* H alleles are the most frequent H alleles in those distantly related species too. Either red *Pundamilia* populations acquired these alleles once or multiple times from other species through introgressive hybridization, or the shared ancestor of red and blue *Pundamilia* possessed all the P and H alleles. In either scenario, the H and P allele split must pre-date the origin of the blue and red *Pundamilia* species.

Selection on the *LWS* gene

We analysed sequences up- and down-stream of *LWS* in a population (Python) that exhibits strong divergence in *LWS* but only weak differentiation at microsatellite loci. Sliding-window F_{ST} analysis revealed at least 6 times greater divergence in the *LWS* gene exons and in 2 kilobases (kb) of upstream sequence ($F_{ST} > 0.8$; Supplementary Fig. 2a) than in the downstream sequences ($F_{ST} < 0.15$), and more than 50 times greater divergence than at microsatellite loci (Table 2). Together with results of McDonald tests³⁸ and HKA tests³⁹ (Supplementary Table 4 and Supplementary Information), this is consistent with a recent selective sweep in the red species, associated with increased presence in a red-shifted environment.

Divergence in the *SWS2A* opsin gene

We sequenced the *SWS2A* opsin gene at two islands to test for divergence at the short-wavelength end of the light spectrum. Out of 10 variable nucleotide positions, 5 were synonymous and 5 were located in introns (Supplementary Table 7). At Kissenda, the *SWS2A* sequences were variable in both phenotypes, and differentiated between them ($F_{ST} 0.1$, $P < 0.01$). At Makobe, a single *SWS2A* sequence variant was almost fixed in *P. pundamilia*, and the species were more strongly differentiated, although not as strongly as in *LWS* (F_{ST} : 0.437, $P < 0.001$; Fig. 2d).

Female mating preferences

Experiments and field data suggest that female *Pundamilia* use male colour as an important mate choice cue^{19,33,34}. Most wild and laboratory-bred Python island females prefer either blue or red males, but laboratory-bred F_1 -hybrid females, most laboratory-bred F_2 -hybrid females and most Luanso females have no preference between blue and red males^{40,41}. Combining published data^{40,41} with previously unpublished data for 11 females from Python island, we find that the frequency distributions of female mating preferences differ between the islands (G -test, $P = 0.02$), roughly resembling those of male nuptial colour (compare Fig. 2f with 2b). The distribution at Luanso (38 females) had a single mode on no preference, and a skew towards blue preference. The distribution at Python (27 females) was bimodal.

We analysed Python island non-hybrid and laboratory-bred F_2 hybrid females to ask whether the *LWS* genotype directly determines mating preference. For non-hybrids and hybrids combined, we observed a significant association between individual *LWS* genotype and mating preference ($\chi^2 = 22$, $df = 10$, $P = 0.03$, 10,000 randomizations). However, this relationship was not significant when restricted to F_2 hybrid females ($\chi^2 = 10.2$, $df = 6$, $P = 0.13$, 10,000 randomizations). Hence, variation in the *SWS2A*-*SWS2B*-*LWS* chromosomal region alone does not strongly predict visual mating preferences in a laboratory environment: some component of mating preference seems independent of it, consistent with biometric estimates that implied that the difference in mating preferences between *P. pundamilia* and *P. nyererei* was due to more than one factor⁴⁰. Modelling light detection, using solar spectrum, water transmission, *Pundamilia* colour patch reflection and *Pundamilia* visual pigment absorption, suggested that a λ_{max} shift of 4 nm towards longer

wavelengths causes a 10% increase in quantum catch for a fish looking at a red patch¹⁶. It seems probable that, in interaction with ambient light in the natural environment, the opsin genotype more strongly determines mating preference than it does under standard laboratory light conditions.

Discussion

Our data on ambient light colour, male nuptial colour, visual pigment λ_{max} and female mating preference indicate sensory drive speciation, which occurred or is maintained by selection without geographical isolation. However, we only observed this under a restricted range of environmental conditions. At all sites with moderately shallow to moderately steep light gradients, two differentiated populations emerged with strong associations between water depth, *LWS* alleles, colouration and preferences (Fig. 3). Strong bimodalities in the quantitative traits colour and preference, strong heterozygote deficiencies at the *LWS* opsin gene, and differentiation at microsatellite loci clearly indicate speciation initiated by strong selection on *LWS*. Very steep light gradients, in contrast, were inhabited by single panmictic populations that showed little variation in *LWS*, even though they contained some variation in colour and mating preference.

The following sensory drive speciation scenario is fully consistent with our data. First, divergent natural selection between light regimes at different water depths acts on *LWS*. Second, sexual selection for conspicuous colouration is also divergent because perceptual biases differ between light regimes. Third, their interaction generates initial deviation from linkage equilibrium between *LWS* and nuptial colour alleles as observed on all but the steepest gradients. Fourth, subsequent disruptive selection due to reduced fitness of genotypes with a mismatch between *LWS* and colour alleles causes speciation, perhaps involving reinforcement-like selection for mating preferences, whereby male nuptial colour may serve as a marker trait for opsin genotype.

The strong association between *LWS* alleles and male nuptial colouration with few or no mismatch genotypes in sympatric species pairs is not restricted to *P. pundamilia* and *P. nyererei* (Table 3, Supplementary Fig. 3 and Supplementary Information). In contrast with these results, we did not find any such discontinuities in the frequency distribution of opsin genotypes along very shallow (between-island) gradients investigated earlier⁹—that is, intermediate *LWS* genotypes predominated in large sections of each gradient. This suggested the presence of divergent selection but the absence of disruptive selection (or the absence of an evolutionary response to disruptive selection). This is consistent with the low migration load predicted from the very small difference in ambient light that migrants between adjacent islands experience (Supplementary Table 1). Despite positive correlations between frequencies of *LWS* alleles and male nuptial colour morphs, and complete fixation of different *LWS* alleles between some populations, speciation as would be indicated first, by strong association between *LWS* and colour and, second, by genotypic and phenotypic discontinuities was not observed on these gradients. This may be due to a difference between the taxa that we studied, but it may also imply that speciation requires disruptive selection, and hence migration and gene flow between habitats^{5,28–30,42}. In contrast, when migration exceeds selection, divergence cannot occur either^{43,44}. This explains the absence of speciation on the steepest of our gradients.

Our results are relevant to conservation because they provide a mechanistic explanation for the collapse of cichlid fish species diversity during the anthropogenic eutrophication of Lake Victoria⁸. Eutrophication changes the slope of environmental light gradients, and, by steepening them, potentially moves sites from the region in parameter space that is permissive of species coexistence into the region that is not. We hope these results help focus attention of biodiversity conservation efforts in Lake Victoria and other lakes to issues of water quality.

Table 3 | LWS opsin allele-group frequency (%) and male nuptial colouration in species of *Pundamilia*

Species	Population	Male nuptial colour type	P	M3	H	Others	n†
<i>P. "Luanso"</i>	Luanso island	Predominantly blue	100	0	0	0	54
<i>P. "Marumbi"</i>	Marumbi island	Predominantly blue	92	0	8	0	24
<i>P. pundamilia</i>	Makobe island	Blue	94	0	0	6	32
<i>P. pundamilia</i>	Igombe island	Blue	83	17	0	0	6
<i>P. pundamilia</i> -like*	Kissenda island	Blue	96	0	4	0	70
<i>P. pundamilia</i> -like*	Python island	Blue	90	1	4	5	82
<i>P. azurea</i> ¹⁶	Ruti island	Blue	100	0	0	0	6
<i>P. nyererei</i> -like*	Kissenda island	Red dorsum	25	0	75	0	52
<i>P. nyererei</i> -like*	Python island	Red dorsum	4	0	96	0	98
<i>P. nyererei</i>	Makobe island	Red dorsum	0	4	96	0	112
<i>P. igneopinnis</i>	Igombe island	Red dorsum	0	0	100	0	6
<i>P. "red head"</i> ¹⁶	Zue island	Red chest	0	0	100	0	6
Total							548

* Hybridizing populations (neither *P. pundamilia* nor *P. nyererei*, but the hybridizing blue (*P. pundamilia*-like) and red (*P. nyererei*-like) populations shown in Fig. 2 (this study)).

† n represents n haplotypes sequenced.

METHODS SUMMARY

Ambient, absorbance and transmittance light spectra were measured with an Ocean Optics PS 1000 spectrophotometer and a 100 µm optical fibre, in the shade between 8:50 and 9:00 in the morning. We calculated the 'transmittance orange ratio' as the ratio of transmittance in the 550–700 nm range over the total visible range. The 'light slope' was obtained by regressing the transmittance orange ratio against distance (m) from the shore along the lake floor. Male fish in breeding colouration were collected by angling and netting; 480 males were photographed immediately in a photo cuvette. Water depth was measured and recorded to the nearest 0.5 m for each of 960 males. To determine the functional relevance of the observed amino acid substitutions in the LWS genes, the sequence of the P allele was reconstructed from the H allele by *in vitro* mutagenesis. The pigments were then expressed, reconstituted and purified as described elsewhere⁹. Absorption spectra of reconstituted pigments were measured before and after irradiation with light (>490 nm). DNA was extracted from fin tissue of 305 individuals and amplified using 11 microsatellite primers. The fragments were analysed on a Beckman Coulter CEQ 8000 Genetic Analysis System. Determination of the opsin genes was as described previously¹³. We sequenced exons 2–5 of LWS (872 bp), which encode the trans-membrane region, from 263 individuals (526 haplotypes). We sequenced exons 1–5 (including introns) of the SWS2A gene from males of Makobe island and Kissenda island. For detection of selection, the LWS gene and its 5-kb upstream and 3.5-kb downstream flanking sequences (total 10.5 kb) were amplified by long PCR⁹ and sequenced from 10 red and 9 blue males from Python island. To determine female mating preferences, we conducted laboratory two-way mate choice assays with females from Luanso island and Python island and laboratory-bred F₂ hybrids from Python island⁴⁰.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 April; accepted 25 July 2008.

- Schluter, D. & Price, T. Honesty, perception and population divergence in sexually selected traits. *Proc. R. Soc. Lond. B* **253**, 117–122 (1993).
- Boughman, J. W. How sensory drive can promote speciation. *Trends Ecol. Evol.* **17**, 571–577 (2002).
- Gray, S. M. & McKinnon, J. S. Linking color polymorphism maintenance and speciation. *Trends Ecol. Evol.* **22**, 71–79 (2007).
- Chunco, A. J., McKinnon, J. S. & Servedio, M. R. Microhabitat variation and sexual selection can maintain male color polymorphisms. *Evolution* **61**, 2504–2515 (2007).
- Kawata, M., Shoji, A., Kawamura, S. & Seehausen, O. A genetically explicit model of speciation by sensory drive within a continuous population in aquatic environments. *BMC Evol. Biol.* **7**, 99 (2007).
- Boughman, J. W. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948 (2001).
- Levring, T. & Fish, G. R. The penetration of light in some tropical East African waters. *Oikos* **7**, 98–109 (1956).
- Seehausen, O., van Alphen, J. J. M. & Witte, F. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808–1811 (1997).
- Terai, Y. *et al.* Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* **4**, 2244–2251 (2006).
- Shichida, Y. *The Retinal Basis of Vision: Visual pigment: photochemistry and molecular evolution* (ed. Toyoda, J.-I.) 23–35 (Elsevier Science, 1999).
- Yokoyama, S., Blow, N. S. & Radlummer, F. B. Molecular evolution of color vision of zebra finch. *Gene* **259**, 17–24 (2000).
- Carleton, K. L. & Kocher, T. D. Cone opsin genes of African cichlid fishes: Tuning spectral sensitivity by differential gene expression. *Mol. Biol. Evol.* **18**, 1540–1550 (2001).
- Terai, Y., Mayer, W. E., Klein, J., Tichy, H. & Okada, N. The effect of selection on a long wavelength-sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. *Proc. Natl Acad. Sci. USA* **99**, 15501–15506 (2002).
- Parry, J. W. L. *et al.* Mix and match color vision: Tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Curr. Biol.* **15**, 1734–1739 (2005).
- Carleton, K. *et al.* Visual sensitivities tuned by heterochronic shifts in opsin gene expression. *BMC Biol.* **6**, 22 (2008).
- Carleton, K. L., Parry, J. W. L., Bowmaker, J. K., Hunt, D. M. & Seehausen, O. Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Mol. Ecol.* **14**, 4341–4353 (2005).
- Spady, T. C. *et al.* Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Mol. Biol. Evol.* **22**, 1412–1422 (2005).
- Genner, M. J. *et al.* Age of cichlids: New dates for ancient lake fish radiations. *Mol. Biol. Evol.* **24**, 1269–1282 (2007).
- Maan, M. E. *et al.* Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. *Proc. R. Soc. Lond. B* **271**, 2445–2452 (2004).
- Endler, J. A. Some general comments on the evolution and design of animal communication systems. *Phil. Trans. R. Soc. Lond. B* **340**, 215–225 (1993).
- Vandermeer, H. J., Anker, G. C. & Barel, C. D. N. Ecomorphology of retinal structures in zooplanktivorous haplochromine cichlids (Pisces) from Lake Victoria. *Environ. Biol. Fishes* **44**, 115–132 (1995).
- Smit, S. A. & Anker, G. C. Photopic sensitivity to red and blue light related to retinal differences in two zooplanktivorous haplochromine species (Teleostei, Cichlidae). *Neth. J. Zool.* **47**, 9–20 (1997).
- Maan, M. E., Hofker, K. D., van Alphen, J. J. M. & Seehausen, O. Sensory drive in cichlid speciation. *Am. Nat.* **167**, 947–954 (2006).
- Endler, J. A. Gene flow and population differentiation. *Science* **179**, 243–250 (1973).
- Schluter, D. & Nagel, L. M. Parallel speciation by natural selection. *Am. Nat.* **146**, 292–301 (1995).
- Nosil, P., Egan, S. R. & Funk, D. J. Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution* **62**, 316–336 (2008).
- Stinchcombe, J. T. & Hoekstra, H. E. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**, 158–170 (2008).
- Doebeli, M. & Dieckmann, U. Speciation along environmental gradients. *Nature* **421**, 259–264 (2003).
- Gavrilets, S. *Fitness Landscapes and the Origin of Species* (Princeton Univ. Press (2004)).
- Leimar, O., Doebeli, M. & Dieckmann, U. Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient. *Evolution* **62**, 807–822 (2008).
- Nosil, P., Vines, T. H. & Funk, D. J. Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59**, 705–719 (2005).
- Seehausen, O. *Lake Victoria Rock Cichlids. Taxonomy, Ecology and Distribution*. (Verduijn Cichlids, 1996).
- Seehausen, O. & van Alphen, J. J. M. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behav. Ecol. Sociobiol.* **42**, 1–8 (1998).
- Stelkens, R. B., Pierotti, M. E. R., Joyce, D. A., Smith, A. M., van der Sluijs, I. & Seehausen, O. Female mating preferences facilitate disruptive sexual selection on male nuptial colouration in hybrid cichlid fish. *Phil. Trans. R. Soc. B* **363**, 2861–2870 (2008).
- Okullo, W. *et al.* Parameterization of the inherent optical properties of Murchison Bay, Lake Victoria. *Appl. Opt.* **46**, 8553–8561 (2007).
- Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–745 (2000).
- Asenjo, A. B., Rim, J. & Oprian, D. D. Molecular determinants of human red/green color discrimination. *Neuron* **12**, 1131–1138 (1994).

38. McDonald, J. H. Improved tests for heterogeneity across a region of DNA sequence in the ratio of polymorphism to divergence. *Mol. Biol. Evol.* **15**, 377–384 (1998).
39. Hudson, R. R., Kreitman, M. & Aguade, M. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**, 153–159 (1987).
40. Haesler, M. P. & Seehausen, O. Inheritance of female mating preference in a sympatric sibling species pair of Lake Victoria cichlids: implications for speciation. *Proc. R. Soc. B* **272**, 237–245 (2005).
41. van der Sluijs, I., van Alphen, J. J. M. & Seehausen, O. Preference polymorphism for coloration but no speciation in a population of Lake Victoria cichlids. *Behav. Ecol.* **19**, 177–183 (2008).
42. Nosil, P., Crespi, B. J. & Sandoval, C. P. Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Proc. R. Soc. Lond. B* **270**, 1911–1918 (2008).
43. Nosil, P. & Crespi, B. J. Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. *Evolution* **58**, 102–112 (2004).
44. Rasanen, K. & Hendry, A. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecol. Lett.* **11**, 624–636 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge the Tanzania Commission for Science & Technology for research permissions, the Tanzania Fisheries Research Institute, and its Muranza Centre director E. F. B. Katunzi, for hospitality and logistical support; M. Kayeba, M. Haluna, S. Mwaiko, M. Haesler and E. Burgerhout for help

with data and fish collection; H. Araki, L. Excoffier, L. Harmon, B. Ibelings, I. Keller, T. Kocher, P. Nosil, M. Pierotti, D. Schluter, A. Sivasundar and O. Svensson for comments on the manuscript; and M. Kawata, J. J. M. van Alphen, K. Young, R. Stelkens and E. Bezault for discussion. This work was supported by Swiss National Science Foundation project 3100A0-106573 (to O.S.), and by the Ministry of Education, Culture, Sports, Science and Technology of Japan (to N.O.).

Author Contributions O.S. conceived and designed the study, collected, photographed and identified fish, measured light and shore slopes, supervised field work, conducted the hybridization experiments, supervised microsatellite analyses and mate choice experiments, and did the statistical data analyses and the writing. Y.T. designed experiments on opsins, did most of the laboratory work and data analysis on opsins, and contributed to writing. I.S.M. collected depth distribution data and did all microsatellite analyses. K.L.C. determined *LWS* sequences from experimental females and contributed to writing. H.D.J.M. collected depth distribution, light data and fish. R.M. determined *LWS* and *SW52A* sequences with Y.T. I.v.d.S. collected fish and conducted mate choice experiments. M.V.S. helped with the microsatellite analysis. M.E.M. collected fish and measured light. H.T. performed analysis of selection pressure with Y.T. H.I. measured opsin pigment absorbance with Y.T. N.O. designed and supervised the laboratory work on opsins and contributed to the writing.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.S. (ole.seehausen@aqua.unibe.ch) or N.O. (nokada@bio.titech.ac.jp).

METHODS

Ambient light gradients and water clarity. Water transparency was measured using a white Secchi disk. Ambient, absorbance and transmittance light spectra between 400 nm and 750 nm were measured every metre between the surface and 3 m water depth with an Ocean Optics PS 1000 spectrophotometer and an optical fibre (100 μ m), using SpectraWin 4.16 software (Avantes). Measurements were taken in the shade, between 8:50 and 9:00 in the morning. We calculated at every depth the 'transmittance orange ratio', which is a property of the water unaffected by variation in solar irradiance, as the ratio of transmittance in the 550–700 nm range (yellow, orange, red) over the total visible range (400–700 nm). The steepness of the light gradient, the 'light slope', was calculated by regressing the transmittance orange ratio against the mean distance (m) from the shore, measured along the lake floor in three transects for every island. The turbidity-mediated between-island light slopes were calculated by regressing the transmittance orange ratio measured at every island at 2 m water depth against the distance (m) from the clear water end of each gradient. The light differential was measured for both types of gradients as the difference between the transmittance orange ratios at the end points of a gradient. The largest possible value is 0.5, which is given when there is no longer any detectable blue light at the deep end of a gradient (transmittance orange ratio = 1 (that is, orange is the only transmitted light); whereas at the surface the full amounts of both blue and orange light are present (that is, transmittance orange ratio = 0.5)).

Frequency and depth distribution of male colouration. Males were collected by angling and gill nets in April and August 2001, February 2003, and January and May 2005. Photos were taken of 11 (Marumbi), 241 (Luanso), 64 (Kissenda), 34 (Python) and 130 (Makobe) males in breeding dress—480 in total—immediately on capture in specially designed photographic cuvettes. Photos were scored on a 5-point (0–4) colour phenotype scale by two to five independent observers, and the mean value was used⁴¹ (Fig. 1). Phenotype scoring of different observers was very similar (Spearman correlations between 0.605 and 0.729, $P < 0.05$). Linear regressions with a quadratic term were fitted to the log-transformed counts of the colour phenotypes from each island separately using R^2 . Frequency distributions were compared between islands by G -tests.

Water depth was measured and recorded to the nearest 0.5 m for each of 960 males. The association between phenotype and water depth was tested for each island separately using ANOVA tests. These males were assigned to colour classes in the field, and only three robust classes were used: blue, intermediate and red (corresponding to classes 0 + 1, 2 and 3 + 4). G -tests were performed to compare depth distributions between islands. The curve-fitting procedure in SPSS (SPSS Inc. 2005) was used to quantify the relationship between strength of association (F -value) and steepness of the light slope.

LWS absorption spectra. *In vitro* mutagenesis of *LWS* for construction of the sequence of P alleles, expression, reconstitution, purification and measurement were performed as described previously⁹ with minor modifications. We measured absorption spectra of reconstituted pigments before and after irradiation with light (>490 nm). On the basis of the λ_{\max} values determined by 3 independent difference spectra calculated from the measurements using independent preparations, we determined the absorption maximum values for each allele with standard errors.

Population genetics of neutral loci. DNA was extracted from fin tissue of 305 individuals (Marumbi 13, Luanso 61, Kissenda 59, Python 84, Makobe 88) and amplified using 11 microsatellite primers developed for these or other

haplochromine species (see Supplementary Methods). We used Arlequin⁴⁶ to calculate observed and expected heterozygosities, to test for significance of departure from Hardy–Weinberg equilibrium for each locus in each population (1 million MCMC permutations), and for significant deviations from linkage equilibrium (10,000 permutations). After sequential Bonferroni correction⁴⁷, 3 out of 55 tests revealed significant deviations from Hardy–Weinberg equilibrium (1 locus each in *P. pundamilia* and *P. nyererei* from Makobe, 1 in *P. pundamilia* from Kissenda), and 2 tests of linkage equilibrium were significant: 1 in *P. pundamilia* from Python island and 1 in *P. pundamilia* from Kissenda island. Because there was no indication of any consistent linkage disequilibrium across populations between any pair of loci, all loci were retained for subsequent analysis. Molecular variance among individuals within and between phenotype groups was visualized in a factorial correspondence analysis performed over individuals in Genetix 4.05 (ref. 48). F_{ST} estimates and their significance were calculated over 100 permutations, as implemented in Arlequin⁴⁶.

Population genetics of opsin genes. Determination of the *LWS* gene was as described previously¹³. We determined the sequences of exons 2–5 of *LWS* (872 bp), which encode the transmembrane region, from 263 individuals (526 haplotypes): Marumbi (12 individuals; 24 haplotypes), Luanso (27; 54), Kissenda (62; 124), Python (90; 180) and Makobe (72; 144). Additionally, we sequenced exons 2–5 of several hundred individuals of other species of Lake Victoria cichlids (Supplementary Fig. 4). Determination of the *SWS2A* gene is described in Supplementary Methods. We sequenced exons 1–5 (including introns) from males of Makobe (16 *P. pundamilia* and 17 *P. nyererei*) and Kissenda (20 blue and 20 red males). F_{ST} values for *LWS* and *SWS2A* sequences were calculated using DnaSP 4.0 (ref. 49). The *SWS2A* sequence (1,930 bp) was split into two putative alleles for the analysis.

Molecular signature of selection on *LWS*. Determination of the *LWS* flanking sequences and the tests for detection of selection were performed as described previously⁹ with minor modifications. The *LWS* gene and its 5 kb upstream and 3.5 kb downstream flanking sequences (total 10.5 kb) were amplified by long PCR⁹ from 10 red and 9 blue males. To reflect the approximate frequencies of *LWS* alleles in the two phenotype populations, we included one heterozygous (H/P) individual of each nuptial colour. The McDonald test³⁸ was calculated with the recombination parameter set to 2, 4, 10, 32 and 1,000 replicates.

Female mating preferences. We conducted laboratory two-way mate choice assays as described elsewhere⁴⁰. Each female was tested in at least 5 trials with 5 different male pairs. A G -test was used to compare the frequency distributions of mating preferences between islands.

45. Venables, W. N. & Ripley, B. D. Modern applied statistics with S. (Springer, 2002).
46. Excoffier, L., Laval, G. & Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50 (2005).
47. Rice, W. R. Analyzing tables of statistical tests. *Evolution* 43, 223–225 (1989).
48. Belkhir K., Borsa P. & Chikhi L., Raufaste, N. & Bonhomme, F. Genetix Version 4.05 for Windows Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France) (1996–2004); <http://www.genetix.univ-montp2.fr/genetix/genetix.htm>.
49. Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., Rozas, R. & Dna, S. P. DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497 (2003).